RNA-seq based expression analysis of the CHO cell protein secretion pathway

Lund, Anne Mathilde; Kaas, Christian Schrøder; Kildegaard, Helene Faustrup; Kristensen, Claus; Andersen, Mikael Rørdam

Publication date:
2013

Document Version
Peer reviewed version

DTU Systems Biology
Department of Systems Biology

RNA-seq based expression analysis of the CHO cell protein secretion pathway

Anne Mathilde Lund[1,*], Christian Schrøder Kaas[2], Helene Faustrup Kildegaard[3], Claus Kristensen[2], Mikael Rørdam Andersen[1]

(1) Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, Kgs. Lyngby, Denmark; (2) Current address: Novo Nordisk, Maaloev, Denmark; (3) Current address: Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hoersholm, Denmark

(*) amalu@bio.dtu.dk

Introduction

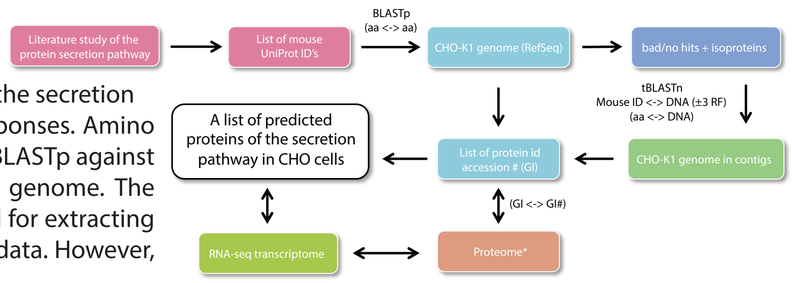
The Chinese hamster ovary (CHO) cell-line is the predominant mammalian industrial cell line being used to produce recombinant therapeutic proteins. Although CHO cells have been used for more than 25 years, the genome sequence was first published in 2011. So far there have been limited studies of the cell biology of the CHO cell and the potential of cell line engineering. To elucidate the poorly understood cellular processes that control and limit recombinant protein production and secretion, a system-wide study was initiated to identify possible engineering targets relevant for therapeutic protein production.

Objectives

- Reconstruction of the complex cellular machineries of the early protein secretion pathway and unfolded protein responses by employing legacy knowledge of mouse
- By using RNA-seq data, a differential gene expression analysis of the constructed CHO secretion pathway would provide a unique possibility for identification of active components to increase the productivity of recombinant proteins.

Strategy of *in silico* resonstruction of CHO cell pathways

First a study was initiated for identifiyng mouse proteins in the litterature associated with the secretion pathway and the ER stress responses. Amino acid sequences were used for BLASTp against the available RefSeq CHO-K1 genome. The identified genes could be used for extracting parts of the mapped RNA-seq data. However, only mapped genes



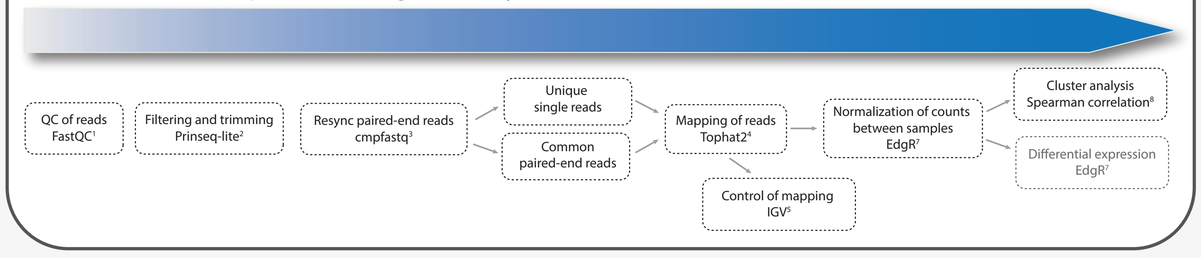
Design and methods for RNA-Seq samples

- RNA was extracted under different growth condidations and treatment from three different CHO cell lines
- Paired-end RNA sequencing was performed by AROS a/s on Illumina HiSeq 2000 platform with a sequencing depth of min. 35 mio reads

Overview of samples

CHO-K1_none		CHO-K1_IgG					CHO-DG44_FVIII		
exp. growth	stationary	exp. growth		stationary		NaBu	exponential growth		
1 samples	1 samples	control	0%NEAA	control	0%NEAA	2 samples	none	medium	high
		1 sample	1 sample	1 sample	1 sample		5 samples	11 samples	2 samples

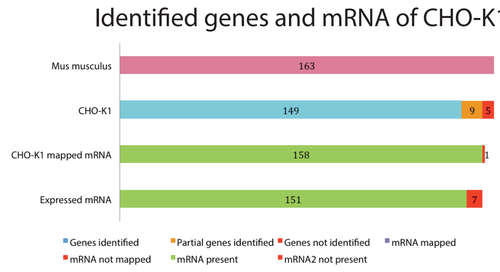
Workflow of RNA-Seq data handling and analysis

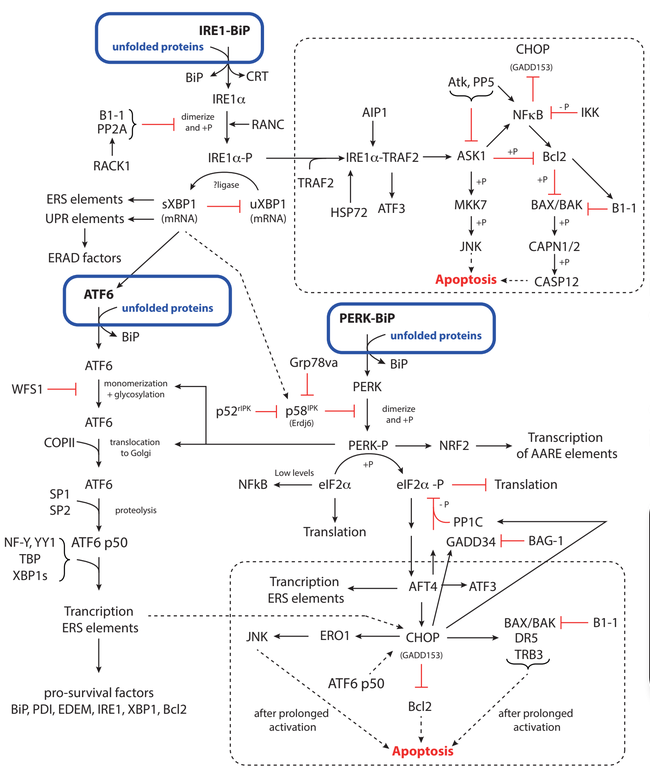


Reconstructed UPR pathway and RNA-Seq data analysis

Proteins associated or linked to UPR pathway were identified by manually curate available litterature on mouse models and cells lines. Furthermore, was the know interactions and dynamics of proteins in the UPR pathway mapped. Found proteins were used to identify CHO-K1 genes of the UPR pathway.

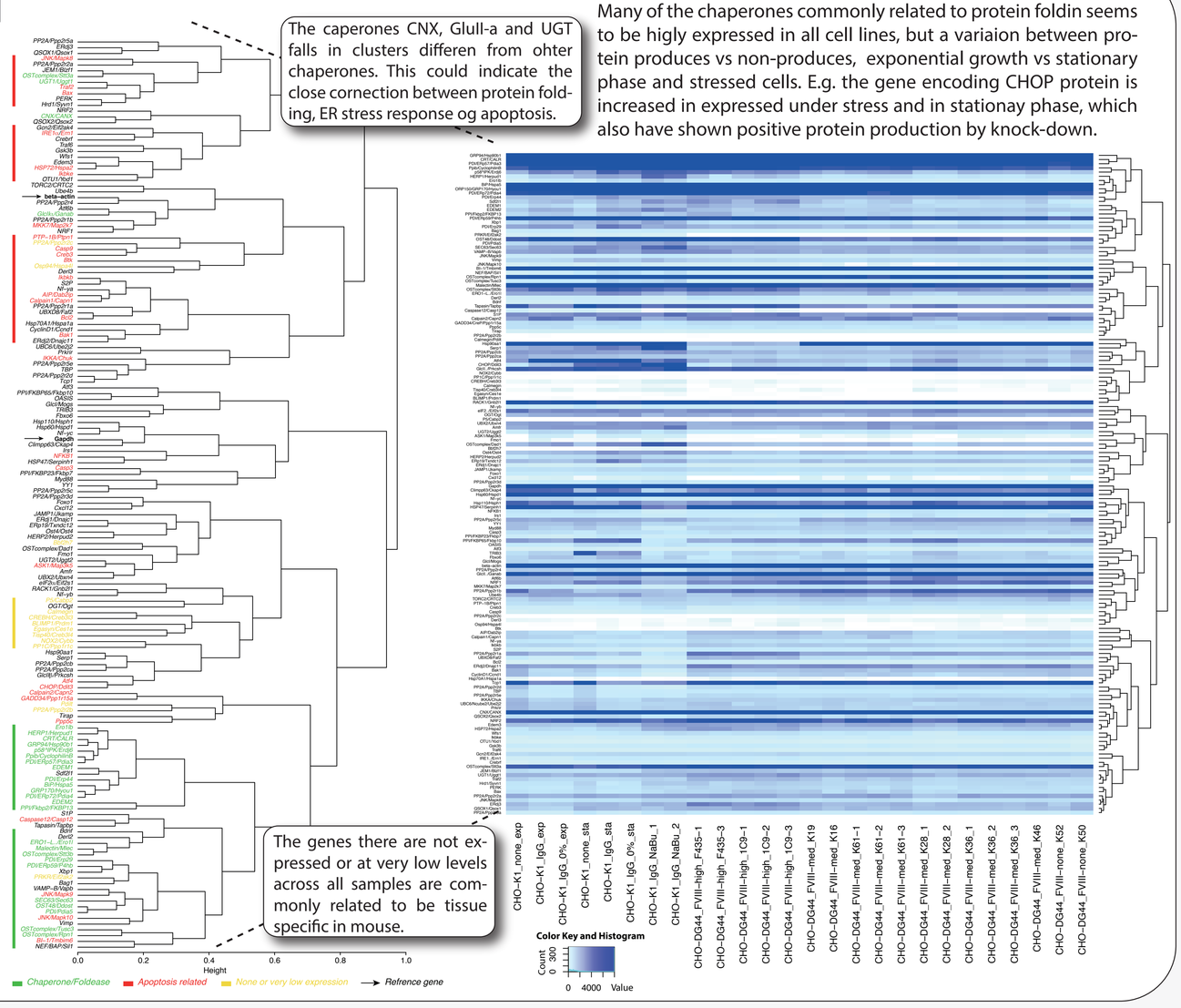
Identified genes and mRNA of CHO-K1





Protein known to be involved in the apoptosic pathway and have a direct interaction is also seen to cluster across the samples in the RNA-Seq data. However, apoptosis can be induced trough different path of the UPR pathway dependend on the cause of stress, which is likely the reason for some of theapoptosis genes to clustered more closely to chaperones and folding proteins

The CHO-K1 genome was sequencing recently and the annotation is at present insufficient, so more than 4 mio. reads was form each sample not possible to map.



Many of the chaperones commonly related to protein foldin seems to be higly expressed in all cell lines, but a variaion between protein produces vs non-produces, exponential growth vs stationary phase and stressed cells. E.g. the gene encoding CHOP protein is increased in expressed under stress and in stationay phase, which also have shown positive protein production by knock-down.

Perspectives

This preliminary study shows the possibilites for using RNA-seq data and cluster analysis to identify new gene clusters based on biological gene expression behaviour that can lead to a greater biological understading of the important industrial cell CHO-K1. n be used for identifying genetic targets for improvement of protein production by overexpression transcription factors og create knock-downs of growth inhibiting.

References

1. B.G. Hansen et al., Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case, Appl Environ Microbiol. (2011) 77(9):3044-51.
2. H.H. Nour-Eldin et al., USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories, Methods Mol. Biol. (2010) 643:185-200.
3. S.S. Tate et al., Secreted alkaline phosphatase: an internal standard for expression of injected mRNAs in the Xenopus oocyte, FASEB J. (1990) 4(2): 227-231.